

Contents lists available at [SciVerse ScienceDirect](#)

# Molecular Phylogenetics and Evolution

journal homepage: [www.elsevier.com/locate/ympev](http://www.elsevier.com/locate/ympev)

## Primate molecular phylogenetics in a genomic era

Nelson Ting\*, Kirstin N. Sterner

Department of Anthropology, 308 Condon Hall, 1218 University of Oregon Eugene, OR 97403, USA

Institute of Ecology and Evolution, 335 Pacific Hall, 5289 University of Oregon Eugene, OR 97403, USA

### ARTICLE INFO

#### Article history:

Available online xxxxx

#### Keywords:

Molecular phylogenetics

Phylogenomics

Multispecies coalescent

Coalescent theory

Morris Goodman

Primate systematics

### ABSTRACT

A primary objective of molecular phylogenetics is to use molecular data to elucidate the evolutionary history of living organisms. Dr. Morris Goodman founded the journal *Molecular Phylogenetics and Evolution* as a forum where scientists could further our knowledge about the tree of life, and he recognized that the inference of species trees is a first and fundamental step to addressing many important evolutionary questions. In particular, Dr. Goodman was interested in obtaining a complete picture of the primate species tree in order to provide an evolutionary context for the study of human adaptations. A number of recent studies use multi-locus datasets to infer well-resolved and well-supported primate phylogenetic trees using consensus approaches (e.g., supermatrices). It is therefore tempting to assume that we have a complete picture of the primate tree, especially above the species level. However, recent theoretical and empirical work in the field of molecular phylogenetics demonstrates that consensus methods might provide a false sense of support at certain nodes. In this brief review we discuss the current state of primate molecular phylogenetics and highlight the importance of exploring the use of coalescent-based analyses that have the potential to better utilize information contained in multi-locus data.

Published by Elsevier Inc.

### 1. Introduction

Fifty years ago, Dr. Morris Goodman published a series of papers investigating the molecular systematics of the living primates (Goodman, 1961, 1962a,b, 1963a,b). These papers were among the very first studies in the field of molecular phylogenetics. Using serological data and protein electrophoresis, he inferred the evolutionary branching order of the lemurs, lorises, galagos, New World monkeys, Old World monkeys, and apes, as well as relationships among certain species within those groups. Furthermore, he proposed a clade-based primate classification that placed humans and African apes in the same family (Hominidae) to the exclusion of orangutans (Pongidae) – a controversial arrangement that met strong resistance at the time [note: this is *prior* to the publication of Willie Hennig's *Phylogenetic Systematics* (1966) in English]. Dr. Goodman's primary motivation for his primate molecular systematic work was to gain a better understanding of our place in the natural world. In honor of his many contributions we thought now would be an appropriate time, on the 50th anniversary of his seminal papers, to assess the field of primate molecular phylogenetics in the age of genomics – in particular the use of consensus vs. coalescent-based approaches when analyzing large multi-locus datasets.

### 2. Analyzing multi-locus data using a single locus framework

One objective of molecular phylogenetics is to provide comprehensive and well-supported phylogenies that reflect species relationships. While this objective has remained constant throughout the last 50 years, the ability to interrogate different types of molecular data and the methodologies used to analyze these data has changed considerably. Earlier studies of molecular evolution inferred species relationships utilizing immunological and hybridization techniques. Although these studies made exciting contributions, particularly with regard to our close genetic relationship to chimpanzees and gorillas (e.g., Goodman, 1961, 1962a,b, 1963a,b), they offered only crude measures of genetic relatedness. Aided by technical advances in molecular biology, later studies were able to overcome this limitation by directly sequencing polypeptide chains and DNA, and from these data infer rates of sequence evolution, gene trees, and divergence times (e.g., as discussed in Goodman, 1996). The traditional molecular phylogenetic framework employed by such studies infers a gene tree and uses it as a hypothesis for the species tree, and infers divergence times across the gene tree through the use of a fossil calibration point and assumption of a molecular clock.

More recently the field of molecular phylogenetics has shifted from single locus studies to multi-locus studies that capture large segments of the nuclear genome. In order to apply the traditional molecular phylogenetic approach to large multi-locus datasets, multiple regions of the genome are concatenated into a supermatrix for analysis (reviewed in de Queiroz and Gatesy, 2007), or

\* Corresponding author at: 308 Condon Hall, Department of Anthropology, 335 Pacific Hall, 5289 University of Oregon Eugene, OR 97403, USA. Fax: +1 541 346 0668.

E-mail address: [nting@uoregon.edu](mailto:nting@uoregon.edu) (N. Ting).

multiple loci are analyzed separately and the resulting trees are combined to form a consensus tree (reviewed in Bininda-Emonds, 2004). While it has been known for some time that gene trees can be incongruent with one another and the species tree (e.g., see, Goodman et al., 1979), these “supermatrix” and “supertree” methods (we will from here on refer to these as “consensus methods”) assume that capturing the most common phylogenetic signal or tree will produce a gene tree that accurately represents species relationships. Within this framework, gene tree incongruence is viewed as an obstacle in the search for true species relationships.

The consensus method approach has been used extensively in recent studies to infer primate phylogenetic relationships from multi-locus data (including Horvath et al., 2008; Wildman et al., 2009; Jameson et al., 2011; Perelman et al., 2011; Wang et al., 2012). For example, Perelman and colleagues (2011) included as many as 191 primate species across 54 genes and 34,941 orthologous base pairs in their study and Jameson and colleagues (2011) sampled eight primate species across more than 1000 genes and one million orthologous base pairs. As a result, we now have a near fully resolved, well-supported, and taxonomically complete genus-level phylogenetic tree for the living primates (Perelman et al., 2011). But is this phylogeny an accurate representation of primate species relationships?

Over the past few years, it has become apparent that consensus methods have theoretical limitations and drawbacks when it comes to the analysis of large multi-locus datasets (Degnan and Rosenberg, 2009; Edwards, 2009). For example, obtaining a false phylogenetic signal through the use of supermatrix methods can occur if substantial amounts of sequence in the dataset are from linked parts of the genome and/or have faster rates of evolution than other parts, thus swamping out potential true phylogenetic signal found at other loci. Results from the supermatrix method are also highly influenced by the selection of alleles in the concatenation process, which is often not taken into account (Weisrock et al., 2012). Furthermore, results from consensus approaches can be confounded by gene tree incongruence due to past hybridization and can fail to uncover such subtleties in the evolutionary history of species. Lastly, consensus methods may lead to very well-supported and well-resolved relationships that are not concordant with the species tree. This can happen at nodes with short branch lengths and/or large ancestral effective population sizes because such nodes are expected to contain gene trees that are frequently discordant from the species tree due to incomplete lineage sorting (reviewed in Degnan and Rosenberg (2009) and Edwards (2009)). In such circumstances, the addition of sequence data in combination with the use of consensus methods could provide an increasingly false sense of confidence in the inferred relationships.

Consensus methods for inferring divergence times across a species tree face similar limitations. Although numerous improvements in inferring divergence times have been developed over the years, including the use of multiple fossil calibration points, allowing for error around calibration points, and incorporating rate heterogeneity across a tree, there remain inconsistencies between the results of primate divergence date studies that are not fully understood. For example, we do not know what characteristics make a sequence dataset ideal for divergence date analyses. Although it stands to reason that longer sequences will produce more accurate dates because they allow for better estimates of mutation rates, obtaining such sequences (i.e., tens of thousands of base pairs) requires concatenating many genomic regions that may not have the same history and/or mutation rates. The implications of this “averaging” approach are not entirely clear as these methods are meant to infer divergence times within a single gene tree. Furthermore consensus methods do not inform us on the timing of actual speciation events because gene tree divergence al-

ways predates species tree divergence by a time period that depends on ancestral effective population size, which consensus methods do not take into account.

### 3. Emergence of a new paradigm in molecular phylogenetics

The theoretical disadvantages of using consensus methods to analyze large multi-locus datasets are driven by a general lack of consideration for the population processes that underlie species relationships. Specifically, by treating the gene tree incongruence caused by various population processes as an artifact to be ignored rather than an important component of species histories, both supermatrix and supertree methods fail to utilize large genomic datasets to their full potential. Continued research into coalescent theory has led to the development of a framework that better realizes this potential for inferring species histories.

Coalescent theory refers to a retrospective model in population genetics that traces all alleles of a gene shared by individuals in a population to a single ancestral copy. The resulting tree-shaped genealogy shows ancestor–descendant relationships among alleles back in time and is known as “the coalescent.” Most commonly attributed to Kingman, coalescent theory was originally described as a set of probability models that demonstrated the time it would take for this process to occur in an idealized population based on genetic drift and population size (Kingman, 1982). After noting the difference between gene trees and species trees in early molecular phylogenetic studies (e.g., Goodman et al., 1979), molecular evolutionary biologists realized that this discrepancy can occur because of the coalescent process (Gillespie and Langley, 1979; Tajima, 1983; Hudson, 1983; Pamilo and Nei, 1988; see also references in Felsenstein, 2004). Through this work, some researchers started to view species trees as statistical distributions of genealogical coalescents, or a “cloud of gene histories” (Wu, 1991; Maddison, 1997; see also references in Felsenstein, 2004). It was also realized that, because coalescent theory is grounded in probability, direct inference of the species tree could be attained by calculating the likelihood that a set of gene tree topologies might exist within a given species tree topology given branch lengths and ancestral population sizes. This theoretical framework is now commonly referred to as the “multi-species coalescent” (Degnan and Rosenberg, 2009). With the collection of multi-locus datasets from non-model organisms becoming more common and an increase in access to computational power, there has been a recent proliferation of programs that use this framework for phylogenetic analysis, each of which implements the theory in different ways (see Liu et al., 2009 and Knowles, 2009 for reviews of some of these programs). In addition to directly inferring the species tree, many of these programs are also capable of inferring population divergence times (as opposed to gene tree divergence times) and estimating ancestral effective population sizes (e.g., BEST and \*BEAST; Liu, 2008 and Heled and Drummond, 2010, respectively). Therefore, in theory these approaches hold a substantial advantage to supermatrix and supertree methods.

#### 3.1. Performance of species tree inference within a coalescent framework

Despite growing support for coalescent-based approaches for species tree inference, their implementation has brought mixed results. In comparison to the supermatrix approach, simulation-based studies have shown that the coalescent approach is more accurate in inferring the species tree when there is a high degree of gene tree incongruence (Edwards et al., 2007; Kubatko and Degnan, 2007). However, empirical studies have shown that coalescent approaches often contain lower levels of support and/or resolution than consensus approaches (e.g., Leache, 2010; Townsend et al., 2011;

Weisrock et al., 2012; Perez et al., 2012). One can thus conclude that either: (1) the coalescent approach is not as powerful as the consensus approach in resolving species relationships, perhaps due to phylogenetic noise from numerous unresolved trees and/or missing data from different partitions (Townsend et al., 2011), or (2) the consensus approach provides a false sense of support, and the multi-species coalescent results are true representations of phylogenetic uncertainty (Leache, 2010; Weisrock et al., 2012).

Currently, it is not entirely clear which of these alternatives is more likely, partly because it is unclear how the coalescent-based approaches perform under a variety of circumstances when handling empirical data. Thus far, studies suggest that taxon sampling, missing data, mutational variance across loci, accurate priors (e.g., mutation rate, effective population size), and the fit of the evolutionary history of the species sampled to the assumptions of the implemented model can all affect the accuracy of multi-species coalescent results (McCormack et al., 2009; Leache and Rannala, 2011; Knowles et al., 2012). However, further research into this area is difficult due to a dizzying number of programs that have been developed, all of which take different parameters into account (e.g., gene tree uncertainty; see Knowles, 2009 and Knowles et al., 2012). Furthermore, numerous empirical studies have found that the most sophisticated Bayesian coalescent methods for species tree inference (those that jointly infer gene trees and the species tree while explicitly modeling individual gene tree uncertainty; e.g., BEST and \*BEAST; Liu, 2008 and Heled and Drummond, 2010) are too computationally intense for some large datasets (Weisrock et al., 2012; Perez et al., 2012). These studies note that chains failed to converge despite running analyses for 200–600 million generations under a variety of priors. This may be the result of very shallow and/or fast divergence events, large ancestral effective population sizes, and missing data, all of which make accurate resolution very difficult (Edwards et al., 2007; McCormack et al., 2009). Lastly, it is likely that there is not one program that implements the multi-species coalescent method in the best possible way, but different programs may be better suited to address different taxonomic groups depending on the underlying species history (Knowles, 2009). Therefore, it is important to minimize missing data, appropriately sample lineages around short internodes, and use programs that employ models that are likely to match the evolutionary history of the species under investigation.

There are many areas for fruitful research and improvement using coalescent-based approaches to molecular phylogenetics. One limitation to most programs has been the assumption that all gene tree incongruence is due solely to incomplete lineage sorting, although some implementations now allow for hybridization (e.g., Meng and Kubatko, 2009; Yu et al., 2011). Another limitation associated with most programs is the assumption that recombination does not occur within loci, which has limited datasets to a relatively small number of short loci that need to be from different regions of the genome. Thus it is unclear how these methods might deal with hundreds, or thousands, or tens of thousands of loci, and long stretches of continuous genomic data cannot be used (see below for a discussion of a potential exception – CoalHMM). These amounts and types of data are quickly becoming available as use of next generation sequencing methods becomes more prevalent in studies of non-model organisms. Obtaining a better understanding of how large datasets, including those with missing data, might be handled is required through empirical testing.

#### 4. Primate molecular phylogenetics in the age of genomics

Primates are an ideal group to use for testing new molecular phylogenetic approaches. Compared to other non-model and

mammalian groups, there is already an abundance of empirical multi-locus data from which to draw. In addition to the Jameson et al., 2011 (1000 + loci) and Perelman et al., 2011 (54 loci) datasets that span the primate order, phylogenomic datasets have been published for targeted primate groups such as the lemurs (18 loci, Horvath et al., 2008), New World monkeys (11 loci, Wildman et al., 2009), colobines (44 loci, Wang et al., 2012), and gibbons (20 loci, Kim et al., 2011).

Nearly all primate molecular phylogenetic studies thus far have used consensus methods for species tree inference. This is a potential concern given that some of the theoretical problems involved in the use of such methods are known to be issues in primates, such as ancient hybridization (see examples in Zinner et al., 2011) and fast speciation events with short internodes (e.g., gibbons, Kim et al., 2011; colobines, Wang et al., 2012). Recently, two studies have been published that use a multi-species coalescent approach to infer primate species trees and both have contributed valuable insight into the use of these methods. Weisrock et al. (2012) analyzed alleles from 12 loci across the lemurs and showed that choice of alleles can have a strong affect on phylogenetic inference, and Perez et al. (2012) reanalyzed the Perelman et al. (2011) dataset in New World monkey genera and highlighted the lower levels of support and resolution present in the coalescent methods and the lack of convergence when larger datasets with missing data are used. This raises the question as to whether or not some of the primate molecular phylogenetic relationships that have been inferred thus far are indeed an accurate representation of the primate species tree. It is thus important to continue re-evaluating primate molecular phylogenetic relationships using a combination of approaches, and results need to be carefully interpreted knowing that consensus methods have their disadvantages and coalescent methods are still being developed and tested. In addition, species tree inference is only one aspect of many multi-species coalescent methods as population divergence times and ancestral population sizes can also be estimated, but this aspect of the framework has been under utilized and needs to be further explored.

Lastly, we would like to note one implementation of coalescent theory in phylogenetics that has been used extensively in primates (hominoids in particular), which is the comparison of long stretches of primate genomic data to investigate the evolutionary processes (mutation, drift, gene flow, selection) most important in shaping patterns of genome variation (see Siepel (2009) for an extensive review). Specifically, a coalescent hidden Markov model (CoalHMM; Duthiel et al., 2009) that incorporates recombination has been implemented. This allows multiple histories to be parsed out from long stretches of continuous sequence data for the inference of levels of incomplete lineage sorting, timing of population divergences, ancestral effective population sizes, and ancestral recombination rates (Patterson et al., 2006; Hobolth et al., 2007; Burgess and Yang, 2008; McVicker et al., 2009; Mailund et al., 2011; Scally et al., 2012.). To our knowledge, this model has not been implemented to infer species relationships and has instead used fixed relationships among the hominoids for the inference of other population and genomic parameters. However, such a model should also enable one to infer the most likely species relationships and could be implemented in whole genome comparisons among multiple primate species, given sufficient computational power. Primates are an ideal study group for testing such whole genome approaches, as there are more primate draft genomes available on Ensembl than any other mammalian group. Continued analysis of primate genomes within a coalescent phylogenomic framework will lead to further insights into species relationships, divergence, adaptation, and evolution (Siepel, 2009).

## 5. Conclusion

It is clear that a paradigm shift has occurred in molecular phylogenetics that blurs the line between population genetics and phylogenetics. However, despite theoretical advantages and a recent proliferation of implementations, there is still much to learn about these coalescent approaches to species tree inference, particularly when applied to real data. Primates are an ideal taxonomic group for empirical testing of newly developed coalescent methods due to the availability of multi-locus datasets and a variety of primate genomes. Continued exploration of new methods with genomic data will ensure that the field moves towards Dr. Goodman's goal of inferring an accurate primate family tree and enabling a more complete understanding of human and nonhuman primate evolution.

## Acknowledgments

We would like to thank Derek Wildman, Ryan Raaum, and two anonymous reviewers for their insightful comments on an earlier version of this manuscript.

## References

- Bininda-Emonds, O.R.P., 2004. The evolution of supertrees. *Trends Ecol. Evol.* 19 (6), 315–322.
- Burgess, R., Yang, Z., 2008. Estimation of hominoid ancestral population sizes under bayesian coalescent models incorporating mutation rate variation and sequencing errors. *Mol. Biol. Evol.* 25, 1979–1994.
- de Queiroz, A., Gatesy, J., 2007. The supermatrix approach to systematics. *Trends Ecol. Evol.* 22 (1), 34–41.
- Degnan, J.H., Rosenberg, N.A., 2009. Gene tree discordance, phylogenetic inference and the multispecies coalescent. *Trends Ecol. Evol.* 24, 332–340.
- Dutheil, J.Y., Ganapathy, G., Hobolth, A., Mailund, T., Uyenoyama, M.K., Schierup, M.H., 2009. Ancestral population genomics: the coalescent hidden Markov model approach. *Genetics* 183 (1), 259–274.
- Edwards, S.V., 2009. Is a new and general theory of molecular systematics emerging? *Evolution* 63, 1–19.
- Edwards, S.V., Liu, L., Pearl, D.K., 2007. High-resolution species trees without concatenation. *Proc. Natl. Acad. Sci. U.S.A.* 104 (14), 5936–5941.
- Felsenstein, J., 2004. *Inferring Phylogenies*. Sinauer Associates, Sunderland, Massachusetts.
- Gillespie, J.H., Langley, C.H., 1979. Are evolutionary rates really variable? *J. Mol. Evol.* 13, 27–34.
- Goodman, M., 1961. The role of immunochemical differences in the phyletic development of human behavior. *Hum. Biol.* 33, 131–162.
- Goodman, M., 1962a. Evolution of the immunologic species specificity of human serum proteins. *Hum. Biol.* 34, 104–150.
- Goodman, M., 1962b. Immunochimistry of the primates and primate evolution. *Ann. N.Y. Acad. Sci.* 102, 219–234.
- Goodman, M., 1963a. Man's place in the phylogeny of the primates as reflected in serum proteins. In: Washburn, S. (Ed.), *Classification and Human Evolution*. Viking Fund Publications in Anthropology, Chicago, pp. 204–230.
- Goodman, M., 1963b. Serological analysis of the systematics of recent hominoids. *Hum. Biol.* 35, 377–436.
- Goodman, M., 1996. Epilogue: a personal account of the origins of a new paradigm. *Mol. Phylogenet. Evol.* 5, 269–285.
- Goodman, M., Czelusniak, J., Moore, G.W., Romero-Herrera, A.E., Matsuda, G., 1979. Fitting the gene lineage into its species lineage, a parsimony strategy illustrated by cladograms constructed from globin sequences. *Syst. Zool.* 28, 132–163.
- Heled, J., Drummond, A.J., 2010. Bayesian inference of species trees from multilocus data. *Mol. Biol. Evol.* 27 (3), 570–580.
- Hennig, W., 1966. *Phylogenetic Systematics*. University of Illinois Press, Urbana, Chicago & London.
- Hobolth, A., Christensen, O.F., Mailund, T., Schierup, M.H., 2007. Genomic relationships and speciation times of human, chimpanzee, and gorilla inferred from a coalescent hidden Markov model. *PLoS Genet.* 3, e7.
- Horvath, J.E., Weisrock, D.W., Embry, S.L., Fiorentino, I., Balhoff, J.P., Kappeler, P., Wray, G.A., Willard, H.F., Yoder, A.D., 2008. Development and application of a phylogenomic toolkit: resolving the evolutionary history of Madagascar's lemurs. *Genome Res.* 18, 489–499.
- Hudson, R.R., 1983. Testing the constant-rate neutral allele model with protein sequence data. *Evolution* 37, 203–217.
- Jameson, N.M., Hou, Z.C., Sterner, K.N., Weckle, A., Goodman, M., Steiper, M.E., Wildman, D.E., 2011. Genomic data reject the hypothesis of a prosimian primate clade. *J. Hum. Evol.* 61, 295–305.
- Kim, S.K., Carbone, L., Becquet, C., Mootnick, A.R., Li, D.J., de Jong, P.J., Wall, J.D., 2011. Patterns of genetic variation within and between gibbon species. *Mol. Biol. Evol.* 28 (8), 2211–2218.
- Kingman, J.F.C., 1982. The coalescent. *Stochastic Process Appl.* 13 (3), 235–248.
- Knowles, L.L., 2009. Estimating species trees: methods of phylogenetic analysis when there is incongruence across genes. *Syst. Biol.* 58, 463–467.
- Knowles, L.L., Lanier, H.C., Klimov, P.B., He, Q., 2012. Full modeling versus summarizing gene-tree uncertainty: method choice and species-tree accuracy. *Mol. Phylogenet. Evol.* <http://dx.doi.org/10.1016/j.ympev.2012.07.004>
- Kubatko, L.S., Degnan, J.H., 2007. Inconsistency of phylogenetic estimates from concatenated data under coalescence. *Syst. Biol.* 56, 17–24.
- Leache, A.D., 2010. Species trees for spiny lizards (genus *Sceloporus*): identifying points of concordance and conflict between nuclear and mitochondrial data. *Mol. Phylogenet. Evol.* 54, 162–171.
- Leache, A.D., Rannala, B., 2011. The accuracy of species tree estimation under simulation: a comparison of methods. *Syst. Biol.* 60 (2), 126–137.
- Liu, L., 2008. BEST: Bayesian estimation of species trees under the coalescent model. *Bioinformatics* 24, 2542–2543.
- Liu, L., Yu, L., Kubatko, L., Pearl, D.K., Edwards, S.V., 2009. Coalescent methods for estimating phylogenetic trees. *Mol. Phylogenet. Evol.* 53, 320–328.
- Maddison, W.P., 1997. Gene trees in species trees. *Syst. Biol.* 46, 523–536.
- Mailund, T., Dutheil, J.Y., Hobolth, A., Lunter, G., Schierup, M.H., 2011. Estimating divergence time and ancestral effective population size of Bornean and Sumatran orangutan subspecies using a coalescent hidden Markov model. *PLoS Genet.* 7, e1001319.
- McCormack, J.E., Huang, H., Knowles, L.L., 2009. Maximum likelihood estimates of species trees: how accuracy of phylogenetic inference depends upon the divergence history and sampling design. *Syst. Biol.* 58, 501–508.
- McVicker, G., Gordon, D., Davis, C., Green, P., 2009. Widespread genomic signatures of natural selection in hominid evolution. *PLoS Genet.* 5.
- Meng, C., Kubatko, L.S., 2009. Detecting hybrid speciation in the presence of incomplete lineage sorting using gene tree incongruence: a model. *Theor. Popul. Biol.* 75, 35–45.
- Pamilo, P., Nei, M., 1988. Relationships between gene trees and species trees. *Mol. Biol. Evol.* 5 (5), 568–583.
- Patterson, N., Richter, D.J., Gnerre, S., Lander, E.S., Reich, D., 2006. Genetic evidence for complex speciation of humans and chimpanzees. *Nature* 441, 1103–1108.
- Perelman, P., Johnson, W.E., Roos, C., Seauanez, H.N., Horvath, J.E., Moreira, M.A., Kessing, B., Pontius, J., Roelke, M., Rumpler, Y., Schneider, M.P., Silva, A., O'Brien, S.J., Bacon-Slattery, J., 2011. A molecular phylogeny of living primates. *PLoS Genet.* 7, e1001342.
- Perez, S.I., Klaczko, J., dos Reis, S.F., 2012. Species tree estimation for a deep phylogenetic divergence in New World monkeys (Primates: Platyrrhini). *Mol. Phylogenet. Evol.* <http://dx.doi.org/10.1016/j.ympev.2012.07.014>
- Scally, A., Dutheil, J.Y., Hillier, L.W., Jordan, G.E., Goodhead, I., Herrero, J., Hobolth, A., Lappalainen, T., Mailund, T., Marques-Bonet, T., McCarthy, S., Montgomery, S.H., Schwalie, P.C., Tang, Y.A., Ward, M.C., Xue, Y., Yngvadottir, B., Alkan, C., Andersen, L.N., Ayub, Q., Ball, E.V., Beal, K., Bradley, B.J., Chen, Y., Clee, C.M., Fitzgerald, S., Graves, T.A., Gu, Y., Heath, P., Heger, A., Karakoc, E., Kolb-Kocinski, A., Laird, G.K., Lunter, G., Meader, S., Mort, M., Mullikin, J.C., Munch, K., O'Connor, T.D., Phillips, A.D., Prado-Martinez, J., Rogers, A.S., Sajjadian, S., Schmidt, D., Shaw, K., Simpson, J.T., Stenson, P.D., Turner, D.J., Vigilant, L., Vilella, A.J., Whitener, W., Zhu, B., Cooper, D.N., de Jong, P., Dermitzakis, E.T., Eichler, E.E., Flicke, P., Goldman, N., Mundy, N.I., Ning, Z., Odom, D.T., Ponting, C.P., Quail, M.A., Ryder, O.A., Searle, S.M., Warren, W.C., Wilson, R.K., Schierup, M.H., Rogers, J., Tyler-Smith, C., Durbin, R., 2012. Insights into hominid evolution from the gorilla genome sequence. *Nature* 483, 169–175.
- Siepel, A., 2009. Phylogenomics of primates and their ancestral populations. *Genome Res.* 19, 1929–1941.
- Tajima, F., 1983. Evolutionary relationships of DNA sequences in finite populations. *Genetics* 105 (2), 437–460.
- Townsend, T.M., Mulcahy, D.G., Noonan, B.P., Sites Jr., J.W., Kuczynski, C.A., Wiens, J.J., Reeder, T.W., 2011. Phylogeny of iguanian lizards inferred from 29 nuclear loci, and a comparison of concatenated and species-tree approaches for an ancient, rapid radiation. *Mol. Phylogenet. Evol.* 61, 363–380.
- Wang, X.-P., Yu, L., Roos, C., Ting, N., Ping, C., Wang, J., Zhang, Y.-P., 2012. Phylogenetic relationships among the colobine monkeys revisited: new insights from analyses of complete mitochondrial genomes and forty-four nuclear non-coding markers. *PLoS One* 7 (4), e36274.
- Weisrock, D.W., Smith, S.D., Chan, L.M., Biebouw, K., Kappeler, P.M., Yoder, A.D., 2012. Concatenation and concordance in the reconstruction of mouse lemur phylogeny: an empirical demonstration of the effect of allele sampling in phylogenetics. *Mol. Biol. Evol.* 29 (6), 1615–1630.
- Wildman, D.E., Jameson, N.M., Opazo, J.C., Yi, S.V., 2009. A fully resolved genus level phylogeny of neotropical primates (Platyrrhini). *Mol. Phylogenet. Evol.* 53, 694–702.
- Wu, C.I., 1991. Inferences of species phylogeny in relation to segregation of ancestral polymorphisms. *Genetics* 127, 429–435.
- Yu, Y., Than, C., Degnan, J.H., Nakhleh, L., 2011. Coalescent histories on phylogenetic networks and detection of hybridization despite incomplete lineage sorting. *Syst. Biol.* 60, 138–149.
- Zinner, D., Arnold, M.L., Roos, C., 2011. The strange blood: natural hybridization in primates. *Evol. Anthropol.* 20, 96–103.